SWEET BRIAR COLLEGE
3 2449 0511123 4

378.05
S9T8b

SWEET BRIAR COLLEGE LIBRARY
SWEET BRIAR COLLEGE

Vol. VII. NOVEMBER, 1923 No. 1

BULLETIN
Sweet Briar College

SWEET BRIAR, VIRGINIA

The Mental Test---A Critical Discussion

BY

LUCY SHEPARD CRAWFORD, Ph. D.

Published by Sweet Briar Institute
November—January—April—June

s second-class matter at the post-office, Sweet Briar, Virginia

ARCHIVES
LD
7251
.S92
B85
1923
v.7,no.1

The Mental Test--A Critical Discussion*

For several years past, the Departments of Education and of Educational Psychology throughout the United States have been intensely interested in the problem of testing human intelligence —of finding an objective measure of mental capacity. The high enthusiasm of the pioneers and of their disciples has now carried the problem beyond the experimental laboratory, until today it has become a practical question which teachers in all departments, in all grades of school from the day-nursery to the university, must sooner or later face and answer for themselves.

As in all questions of educational theory and practice, it may be dangerous to accept blindly and without question or limitation, the judgment of the enthusiast or the expert, no matter how firmly that judgment may seem to be based upon statistics and upon interpretations deduced from practical experience. Especially when a theory or practice is so revolutionary in character, and in such obvious opposition to traditional practice— especially then, are teachers under obligation to scrutinize it with the utmost care, in an effort to understand the fundamental principles upon which the theory is based, as well as the probable or possible effects which that theory seems likely to produce when put into general practice.

As yet, it seems fair to say, opposition to mental testing has been largely unorganized, and inarticulate. The great preponderance of testimony and opinion voiced at teachers' conferences and in the educational journals, has been in favor of the tests. And the army of professional adherents seems to be increasing both in enthusiasm and in number. It is nevertheless indisputable that a considerable number of professional teachers are either in active opposition, as individuals, or have assumed an attitude of indolent neutrality. With the neutrals we are not now immediately concerned, except to suggest that it behooves them to reach a decision without further delay. Even among the neutrals, however, there are many who naturally rebel against the intelligence test in particular cases, and refuse to accept the results when those results conflict with their personal judgment.

---

*A paper read before the Twenty-eighth Annual Meeting of the Association of Colleges and Secondary Schools of the Southern States, held at Richmond, Virginia, December 7, 1923.

To those who challenge the value of the intelligence test as an educational instrument, and as a measuring-rod of innate capacity, or as a reliable means of prognosticating success in school or in life, we now turn our attention. As in all innovations, the burden of proof rests with the mental testers. The claim is made that the new type of test measures innate mental capacity; and that it is not only an invaluable supplement to the old type 'achievement' examination, but may be, in time at least, an adequate substitute as one of the controlling factors in determining admission to college, in promotion from grade to grade, in award of scholarships, etc. The claim is further made that the intelligence test is the only 'objective' and 'scientific', and therefore the only adequate and reliable, method we have of classifying children according to their mental endowment, and of prophesying probable success in school work, or in other occupations of life.

Have the champions of the intelligence test made good their claims, either in practice or in principle? This is what the critics emphatically deny. The critics maintain: (1) A student receiving a low score on an intelligence test may actually possess superior power of achievement, including both intellectual power and power based on character. (2) Conversely, a high rating in an intelligence test is no sure indication that the individual has superior mental power, or is a truly superior individual when considered as a whole.

The challenge on the part of the critics is not confined merely to questioning the validity of results obtained by particular forms of the tests; the more serious challenge involves the fundamental hypothesis upon which all forms of mental tests are based; namely, the possibility of measuring human intelligence, either individually or collectively, by an objective, mechanical, standardized test, in such a way as to foretell mental capacity in advance of achievement. In other words, criticism is leveled against the mental test not only in practice but also in theory.

For the sake of clearness, it may be well to consider separately these two aspects of the question—the practical and the theoretical—in so far as it is possible to make a sharp division.

I. OBJECTIONS BASED UPON PRESENT PRACTICE.

First, what are some of the limitations and difficulties implicit in the tests as now administered—limitations and difficulties which may or may not be inherent, but which, nevertheless, do vitiate the value and accuracy of the results obtained? We might consider these criticisms of detail under the following headings: (1) Type of question; (2) Speed-test; (3) Endurance-test; (4) Nervousness; (5) Reaction of Student and Parent; (6) Relation between Teacher and Student; (7) Correlation of Results with 'Academic Success'; (8) Objectivity of the Mental Test.

1. Type of Question.

Students themselves often bring the charge that the tests include so many "catch" and "tricky" questions that time is lost in answering each question in the endeavor to avoid falling into a possible trap. The further criticism is brought that many of the questions deal with trivialities, such as the identification of characters famous in the motion-picture world, or the identification of advertisement-catchwords, or trade-marks employed by chewing-gum or tobacco manufacturers, or questions referring to professional baseball, or to the manufacture of automobiles, to dress-fashions, tooth paste, etc. Correct answers to such questions are based upon the observation of the "passing show", where lack of observation may easily be a mark of superior intelligence, a mark of good breeding, and more careful rearing, and an indication that the child is interested in more important things, and has a deeper appreciation of the higher values of life.

Another type of question propounds a problem, offers a number of possible answers, and requires the student to check one of these answers as correct. No comment or qualification is permitted. One example will suffice to illustrate the artificiality of this type of question. Problem presented for solution: "If it rains when you are starting to go for a doctor, what should you do?" The student is instructed to choose between the following alternatives: (1) stay at home; (2) take an umbrella; (3) wait till it stops raining, or (4) take a bath. Presumably the "correct" answer is "Take an umbrella." The thoughtful student might quite reasonably hesitate to choose any one of these alternatives. His answer would depend on the purpose of his prospective visit to the doctor. If the need for the doctor were very urgent, the student might feel that he should hurry off without stopping for an umbrella. If the need were not urgent, it might be the part of wisdom to "stay at home", or "wait till it stops raining". The fourth alternative could hardly fail to be distracting in its absurdity. Is it a mark of a "superior" mind to check the second alternative as the answer that is probably expected or as the most likely of the four possibilities? Or, is the more critical and analytical power a mark of superiority?

A similar type of question makes a statement, requires the student to answer categorically "yes" or "no", as to its truth or falsity. For instance: "Cleopatra loved Mark Antony." What thoughtful student could stamp that statement either "true" or "false", without qualification? It is maintained by the critics that this "either-or", "yes-no" type of question tends to present for solution abstract, artificial problems, and fails to test ability to meet concrete situations or solve real problems. The artificiality of such questions and their limitations become apparent when we consider how seldom we can meet adequately any concrete situation by answering "yes" or "no" flatly and categorically, and how often choosing between two or more stipulated alternatives means choosing between evils, whereas the real solution involves the consideration of other conditions.

An interesting, though indirect and unconscious, comment on this type of question is to be found in "*A Referendum of Psychologists—A Survey of Opinion on the Mental Tests,*" by F. M. Freeman, Century Magazine, December 1923 (pages 237-245). In this referendum, Mr. Freeman submitted to a number of psychologists various statements referring to the intelligence tests, asking for assent or dissent in each case. In reporting the replies, Mr. Freeman remarks that "the various persons . . . probably do not, in all cases, wish to be considered as agreeing in rigid fashion to the precise forms of statements which were presented to them." (page 238) As one of the psychologists, included in the referendum, expressed it: "In almost every case . . . I should want to make certain qualifications and interpretations, so that I hope you will not make any rigid use of these replies of mine." (page 238)

In general, the questions fail to test the power of *constructive* thought. They are so framed that the student may write his answer in the simplest, most mechanical way—sometimes by a check-mark, or underscoring a word, or writing a figure. There is no *formulation* of thought on the part of the student, no demand made upon the student to express his thought in his own words. Furthermore, due to the type of question and the pressure of time, students often, perhaps quite generally, skip those questions that puzzle them, and rush on to those that can be answered "on the spur of the minute", with the minimum of thought. They seem to proceed on the theory that the sins of omission are not so serious as the sins of commission, and that it is better to leave a blank than to run the risk of making a mistake. Such a procedure on the part of the student is recognized and deplored by the mental tester. How far this element invalidates the accuracy of statistics in mental testing is still an open question.

2. Speed-test.

On the one hand, we have the speed-test, which places a premium on quickness of response, and penalizes the mind that works more slowly, more deliberately, but often more surely. Such a test over-emphasizes the importance of what might be called the power to make a "snap-judgment", and under-estimates the value of the power of reflection and contemplation. And yet, especially in the higher reaches of intellectual endeavor, the power of reflection is the more to be prized. It may be possible, no doubt, for the naturally quick mind to be tempered and disciplined so as to acquire the power of reflection; just as it may be possible for a person with a slower mind, through training and experience, to quicken his mental tempo. In both cases, however, such training comes only with the process of time and by continued effort, and rare indeed is the individual where the combination has been effected at the age when intelligence tests are usually administered.

3. Endurance-test.

On the other hand, effort has been made to eliminate the difficulty of the speed-element. For instance, in two of the three sections of the Thorndike College Entrance Intelligence Examination, used by Columbia College, it is maintained that the speed element is minimized, and the *endurance* of the student severely tested.* In this examination, within a period of two hours and fifty minutes,† the candidate is called upon to perform twenty-one different kinds of task, including "no less than 460 different items of graded complexity."‡ We would all agree, probably, that that is a test of endurance!

The candidate knows the length of time at his disposal, he is familiar with the type of examination and may know the number of questions to be answered, although obviously it is not possible for him to review the particular questions as a whole before beginning the test. It is reasonable to suppose that throughout the two hours and fifty minutes, the candidate is working under pressure, and at high tension. The speed-element is, therefore, not successfully eliminated. Faced with questions in such large number, and in such variety, the student is forced to pass rapidly from one to another, in a manner that suggests the childish game of "hop-skip-and-jump". It may be claimed that one of the most important functions of the new examination is to test this very agility. The value of the intelligence test as a mental gymnastic may, however, be seriously questioned when that mental gymnastic consists largely, if not entirely, of *jumping at conclusions.*

Returning to the question of endurance: Many of us from our own experience know that a three-hour 'content' examination is so severe a strain both mentally and physically that it often places the truly superior student at a serious disadvantage. In the examination of the old type, however, all the questions may be viewed as a whole, and the allotted time divided at the discretion of the individual student so as to meet his own peculiar needs. In such a case, the demand made upon the individual's

---

**Measurement in Higher Education,* by B. D. Wood, pages 26, 29, 30, *et passim.*

†Note the "scientific accuracy" of limiting the time, not to the round three hours, or two hours and a half, but precisely to 2 hrs. 50 min.! A similar instance of precision is to be found in *Measuring Minds,* by Caroline E. Myers and Garry C. Myers, under "General Directions to Examiners": "Time should be recorded with great precision, exactly to the second, and from the word, 'Go'. A stop watch is essential." (page 46)

‡*Measurement in Higher Education,* page 30.

poise and self-control is not nearly so great as when the problems to be faced are unknown quantities and come under view bit by bit, page by page. Very often this working under pressure would result in more serious mental and physical exhaustion in the case of the superior mind than in the case of the inferior mind, because a high degree of mentality is so often accompanied by a smaller amount of physical endurance, together with a high degree of sensitivity, or with what we often call the "artistic temperament", which, however valuable it may be in many respects, is easily thrown off its balance. By "artistic temperament" I mean in general the more highly sensitized emotional nature, which should be cherished as a precious spiritual leaven in any academic community. Moreover, the endurance-test may be quite as disastrous to the quick mind as to the slower and more deliberate type of mind.

It may be argued that poise and self-control are valuable assets in the field of intellectual endeavor, as in other fields, and that lack of these qualities may spell failure in college. Very true, but it is equally true that they represent only two of many desirable qualities that go to make up the successful student. Furthermore, in their very nature, poise and self-control result from the training and discipline which come only with time and experience. They are not innate, nor may they be acquired by many gifted individuals except at the cost of long-continued and cumulative effort. Their development should be gradual, and seems to be a legitimate function of the college, as well as of the elementary and secondary schools. In other words, it seems hardly justifiable to debar a student from college, or to rank his mentality low, for lack of a high degree of self-control, especially since that very lack may be an indication of a high degree of sensitivity and a deeper appreciation of the intangible beauties of life.

4. Nervousness.

The nervous strain involved in both the speed-test and the endurance-test is further intensified by the consciousness on the part of the student that this examination is not calculated to test his knowledge of any particular subject or stipulated group of subjects, in which he is more or less well prepared, and in which he may reasonably expect to give an account of himself in direct proportion to the effort he has previously expended in study. On the contrary, this new examination is primarily to measure his innate mental capacity—something for which the individual is in no way responsible.

To the serious student, infinitely more is at stake in the one case than in the other. If he fails or receives a low grade in the old-type examination, there is always the opportunity of retrieving lost ground by more assiduous application, or of regaining his self-respect by excelling in other subjects more suited to his particular abilities. But in this new type of examination, there is a finality that is appalling. If the student fails or receives a low grade in the intelligence test that seems to be final. It means that his mental equipment is below par. It means that no amount of effort on his part can remedy that defect. He will be a "marked man" for the rest of his college course; he will be under a disadvantage in his relations with his instructors. He may even fall in his own estimation and come to the conclusion that he is doomed to mediocrity. The more serious and ambitious the student, the more likely is he to be affected by these considerations and, if inclined to nervousness, the less likely is he to give a creditable account of himself.

The possibility and danger of nervousness on the part of the candidate are frankly recognized by the supporters of the intelligence test. It is admitted that nervousness defeats the aim of the test, renders the individual's score unreliable as an index of his true capacity; likewise, the inaccuracy of individual scores nullifies the anticipated accuracy in the average group scoring. Here again, we find ground for questioning the reliability of statistics published in relation to intelligence tests. Effort is made on the part of examiners to allay nervousness in various ways, and to discount the results when nervousness is observed. The difficulties along this path are many: 1) It would be extremely difficult, if in any case possible, to allay nervousness which had arisen from the causes noted above. 2) The larger the group to be examined (from 400 to 500 at the Columbia September examination), the more infectious would panic become, and the more difficult would it be to detect individual nervousness. The presence of a proctor for every fifty students, as at Columbia, might tend to increase rather than to decrease the probability of panic. 3) Even in small groups, the outward composure of a student may effectively conceal inner excitement, especially when the examiner and student are unknown to each other.

5. Reaction of Student and Parent.

What are some of the possible reactions of the student to a scoring on the basis of a mental test? Here we have to note two different methods of procedure. In the one case, the student may be informed, directly or indirectly, of his rating. In the other case, this information is withheld from him, even though he may seek to know.

This difference of procedure, itself, involves a serious question: *Is the teacher, or school administration, justified in withholding such information from the mature student, or from the parents of the immature student?* So long as intelligence tests were in the experimental stage, it might have been defensible for the experimenter to keep this information to himself, and to look upon the individual as suitable "laboratory material" to be used for the "advancement of science". But, just as soon as the results of the intelligence test are considered definitive in a particular case, and are to be used as an index of a student's ability to grade him in school, or to grant or refuse him admission to college, or for any other practical purpose, then I maintain, the student or his natural guardian has a right to such information.

The more serious the student the more keen will be his interest in the result of such a test, and the more uneasy will he feel if he is not told his score. He is very apt to feel that such a rating is much more significant than a mark in a 'content' examination. It is a distinctly personal matter, and it will be very difficult for him to understand the justice of keeping him in ignorance. If school authorities are successful in discouraging this natural interest on the part of the student, their success is measured by the degree to which they have suppressed the student's intellectual curiosity—that fundamental spirit of inquiry, which since the days of Plato has been considered the golden key to learning.

Moreover, if the student is not told his score, he is bound to speculate as to the probable results. His conjectures may lead him far astray. Not knowing how the various questions are 'weighted', or the relation between speed and accuracy, between quantity and quality, the superior student might easily be led to suppose he had given a poor account of himself (even when he had not), and in a similar condition of uncertainty, the poor student might optimistically conclude that he had received a high rating. To be sure, it is the practice of some teachers not to tell students their marks in the traditional examination, even at the end of the semester. The difficulty in such a case is not so serious because, as already suggested, the traditional examination is not so all-inclusive or final as the mental test, and moreover, it is possible for a student later to check up his answers in a 'content' examination, and reckon his grade with reasonable accuracy.

As matter of fact, even when the general rule prevents official notification, particular students very often are informed of their respective scores. Let us now consider some of the possible results when the student knows he has received a rating of "mediocre" or "superior". In proportion as the tests are considered definitive, the effects may be serious.

A student, placed in the lower grade, may easily develop an 'inferiority complex', and a painful self-consciousness. Even if this does not lead to a morbid frame of mind, it may easily lead to indifference and indolence. Suppose a student accepts as final the judgment of the test that he has only mediocre capacity. So long as he keeps his work on the level of mediocrity, he is quite satisfied with himself and has the self-complacent feeling that no one has the right to expect anything more of him. He is doing the best he can. In other words, the result of the test has robbed him of all incentive to improve the quality of his work. Those who accept the verdict of the mental test as final may also say there is no use in his making an effort, that it is a kindness to let the student know his particular limitations. The answer to that is that even though the effort fail to produce mental improvement or development, the making of the effort *per se* is of value in character-building, and therefore of tremendous educational value. It is often the mediocre student who derives more real benefit from college than the student of superior mind. And who will deny that the conscientious though mediocre student may make a better citizen than the student of superior mind who has glided easily through course after course with the minimum of effort?

Just here is the danger for the student whose mental capacity has been rated "superior". Amidst the many distractions of the modern high school and the modern college, the student who is conscious of his superior mental endowment may easily assume the attitude that he has no need to toil like his less fortunate fellows—his natural endowment will carry him through, he will trust to his wits.

Obviously, under the Old Régime also, the 'inferiority complex' or the 'superiority complex' is likely to be found occasionally within any group of students, but without the all-inclusive verdict of the capacity-test, the inferior student may feel power within him to excel in one or more directions, while falling below par in others; and likewise, the superior student may have the humbling consciousness of weakness in one or more directions. On the other hand, the student is rare who, in the natural course of events, will pass a verdict upon himself of hopeless and unmitigated mediocrity; whereas only the unfortunate few hold such an exalted opinion of themselves that they are ready to class themselves superior in all respects. But when in either case, judgment is the result of an 'objective' method of measurement, the temptation to accept the verdict is almost irresistible.

In this connection, we must also consider the possible effect upon the parents of high or low scores received by their children. Even though it be maintained that it is not expedient or necessary to advise students of their individual scores, it can hardly be maintained that such knowledge can, with justice and wisdom, be withheld from parents, especially when such strenuous and earnest efforts are now being made to enlist the parents' intelligent and active interest in the school-life of the child.

Grades or 'marks' now sent to parents may be a source of pride or disappointment. But when an 'intelligence score' is reported, it purports to be a measure of innate mental capacity, of congenital strength or weakness. A low score, therefore, reflects not only upon the individual child, but upon the whole family; it may even suggest an hereditary taint. Family pride is immediately aroused. Rare indeed is the parent who will not resent any such stigma, or any such imputation. Such a situation would be very likely to result in antagonism between the home and the school—the very thing we are so anxious to avoid.

It seems hardly necessary to labor this point, except to suggest a possible result, in case the report sent home ranks the child in the superior class. There seems to be a tendency already on the part of many parents to minimize the importance of curricular activities in comparison with extra-curricular activities. This tendency might be lamentably increased by the assurance that Sally can do quite as well as her duller playmate, with much less effort. Why then should Sally be encouraged to study? In the creation of such fixed *strata of intelligence,* is there no danger of encouraging "intellectual snobbishness", and "intellectual caste"?

6. Relation between Teacher and Student.

We now turn to another important question: How may a high or a low score in an intelligence test of a particular student affect the teacher's relations with that student? It is often urged that a high score may be used by the teacher as a lever to raise a student to a higher level of achievement. By an appeal to a student's natural pride in a high rating, the student may be prevailed upon to make his work more commensurate with his native capacity. Unfortunately, mere consciousness of capacity is not in itself necessarily stimulating. It may be quite the reverse. Such a superior student may quite sincerely prefer to bend his energies in other directions, and argue quite logically that, as his mental capacity is already on a high level, it behooves him to give more attention to development along other lines, such as sociability, dramatics, rhythmic dancing, team work, etc.

Moreover, is there no danger in such a direct appeal to personal pride? I think there is. When the source of personal pride is merely innate endowment, and not the result of the student's personal effort, such a stimulus may be positively harmful. As matter of fact, of course, this method of stimulation was not originated by the mental testers. It dates back at least to Socrates and Plato, to trace it no further. Furthermore, it is a method which we all use to a greater or less degree. But when we assure a student that he is not working to the best of his ability, we support our suggestion either by reference to his own past record, or we express our personal judgment as to his potential capacity. Under such circumstances, we are apt to appeal to the student to confirm our estimate of his ability. There is always the possibility of mistake in judgment; there is no such certainty or finality as results from the professed 'objectivity' of the intelligence test. The result is, when it is merely a personal matter between teacher and student, the student may leave the conference with a feeling more of humility than of pride.

Turning now to the relation between the teacher and the student who has received a low score in the intelligence test: From time immemorial, one of the most important functions of the teacher has been to stimulate the student to outdo himself, to do better tomorrow than he did today. How can a teacher perform this very important function conscientiously and with enthusiasm, if he believes the student has already reached the limits of his capacity, of a capacity that is fixed and immutable?

At this point it may be pertinent to inquire whether it is not really safer after all to trust to an information test. In the case of the student who gives a good account of himself in a 'content' examination, it is reasonable to assume that the acquisition of that knowledge presupposes a considerable amount of capacity. Moreover, from the point of view of practical classroom instruction, the 'content' examination assembles groups of students with at least a quasi-homogeneous foundation. To be sure, in the modern classroom there seems to be very little information that a teacher has a *right* to suppose the students possess! There seems to be danger, however, that this new type of examination may tend not to mitigate but rather to aggravate a situation that is already serious.

Undoubtedly there is much room for improvement in many 'content' examinations. But at their best, they test not primarily the ability of the student to convert himself into a storehouse of particular facts. On the contrary, they aim to test the student's power of analysis, of expression, of interpretation, of judgment, of reflection and independent thought. The fact that this aim seems to be attained only in rare cases should stimulate teachers to greater effort toward improvement, but should not necessarily lead them to condemn the old-type examination in principle.

In two other important and obvious ways, the student may gain benefit from the 'content' examination, which does not come to him from the intelligence test: Prior to the examination, the student is encouraged to review the subject as a whole, to gain a synoptic view which may be of immense value. The examination itself may prove to be a stimulating mental exercise—not of the "hop-skip-and-jump" variety already referred to, but an exercise in constructive thinking, in looking at a problem from all sides, in formulating answers clearly and coherently. Such an internal discussion on the part of the student should lead to a clarification of thought and expression, which may result in the student's actually knowing more upon leaving the examination than he knew before. Is not this the real test of the educational value of any examnation? You may say this is an ideal that is never to be realized. "Never" is a long time. We might cite instances where it had already been realized. In any event, it is an ideal worthy of pursuit.

The tendency has been perhaps to outlaw the old plan of admission to college, as an utter failure, and as utterly inadequate to cope with the present-day problem of large numbers and of heterogeneous early training. Can we be sure that the old plan has failed in principle, or has fallen short only in practice? Is there any wisdom in the suggestion that we bend our energies toward improvement and development of the old plan, and look for progress through evolution rather than through revolution? Are we quite sure, in other words, that we have consistently realized all the advantages which are potentially latent in the old-type examination?*

It may also be remarked in passing that by focusing attention so sharply on the examination-problem, the mental tester has rendered a valuable servce to the schools. The Muse of History may record that the irritation caused by the intelligence test marks a stage of progress in education, in so far as it has stimulated "reform from within". Certainly the innovators have succeeded in arousing active discontent with the 'content' examination at its worst. That discontent, we hope, will bring us one step nearer a true solution.

---

*As matter of fact, an effort is now being made to combine the old type 'content' examination with the new type 'capacity' test. This is a comparatively recent development, and it is perhaps too early to discuss its possibilities within the time at our disposal. It may, however, be questioned whether such a combination, or compromise, will ever prove satisfactory or practically feasible.

In this same connection, it is also to be noted that the use of the intelligence test in many secondary schools has resulted in the formation of graded sections. That the principle of the graded section has much to commend it, both in secondary school and in college, is gladly conceded, and in so far as the use of the intelligence test has hastened the adoption of this principle, by overcoming administrative objections and practical difficulties, we acknowledge our gratitude. The fact remains, however, that the more enlightened members of our profession have long been urging the advisability of graded sections, and the feasibility of making such a separation on the basis of the previous work of the pupil, plus the personal judgment of the teacher. Furthermore, the danger of making such a division on the basis of the intelligence test as the only, or as the controlling, basis of judgment, may be inferred from our previous discussion.

7. Correlation of Results with 'Academic Success'.

A very persistent argument offered in support of the mental test, as a means of predicting school progress, is the 'high correlation' between the 'intelligence score' and 'academic success'. ('Academic success' is measured in terms of average grades received in college work.) Statistical studies have been made with a view to discovering this correlation and expressing it in numerical terms. If later work in college corresponded exactly with the verdict of the intelligence test, the correlation would be unity (1.00). Columbia College, after a three-year experiment, reports 'correlation coefficients' ranging from .60 to .67.* This result is considered highly satisfactory. A correlation ranging from .40 to .50 is considered worth reporting; anything above .50, worthy of special note. How many unreported correlations

---

*In *Measurement in Higher Education,* page 73-74, Mr. Wood suggests that .84 (rather than 1.00) is the maximum correlation coefficient which can reasonably be expected.

fall below .40 it is impossible to estimate.* The statistical evidence is, therefore, incomplete, and does not offer reasonable warrant for generalization. Moreover, the Columbia experiment, where a relatively high correlation was obtained (as reported in *Measurement in Higher Education*, by B. D. Wood) was conducted under unusual circumstances, with a group of students selected with unusual care, and it covers a period of only three years. Its results therefore can hardly be considered typical or conclusive.

In any case, what value are we justified in attaching to these correlation figures? Our answer hinges on the reliability of statistics. We may say that "figures do not lie", but correlation statistics are based upon the interpretation of figures, and therein lies the possibility of grave error. How far are we ready to trust statistics and averages when we are dealing with the complexities and variability of human beings? Is it not true that in group statistics the individual is submerged? Figures, and especially average figures, are abstractions and cannot therefore give us a reliable measure of the concrete individual. Under the old system, 'marks' are not considered absolutely definitive, nor are they the only criterion of 'academic success'. For many years there has been a happy tendency in our colleges and secondary schools to break away from the fixity and pseudo-accuracy of numerical grading, because there has been a growing feeling that the student must be considered as a whole, and that the numerical method is too fixed and rigid for estimating human values. Is the attempt to express by a mathematical formula the relation between mental capacity and 'academic success' really a step forward or backward?

The validity of the correlation statistics may be further questioned when we consider that very often the teachers who are to be the judges of 'academic success', or failure, know at the beginning of the semester the intelligence scores of the students whose work they are to judge. The more importance the teacher attaches to the intelligence test, the greater will be the tendency to grade the student's later work in accordance with that initial test. If the student himself knows the score, that knowledge may likewise tend to create a spurious correlation between his college work and his real capacity. As has already been suggested, this is especially true when the student receives a low intelligence score.

---

*It is to be noted that the correlations obtained by the same test given to different groups are by no means uniform. Is this an indication of lack of 'objectivity'? Cf. later discussion of the tests as an "objective scale of measurement."

In general, it may be recalled that the statistical method, in its essence, is the method of induction, and shares all the dangers and weaknesses of that method. It is submitted that the statistical evidence, offered in support of the mental tests, does not warrant conclusions drawn because 1) it does not extend over a sufficient length of time; 2) it does not include a sufficiently large number of individuals; 3) it does not present any reasonable degree of uniformity; 4) it does not take sufficient account of "negative instances" and conflicting evidence; 5) it is of such a nature as to be of very questionable value in dealing with the concrete individual.

8. Objectivity of the Mental Test.

Does the intelligence test actually afford a scale of measurement which may, to any reasonable degree, be classed as 'objective'? A recent discussion of the subject defines an objective scale of measurement as follows: "What science means by a perfectly objective scale is a scale in respect to whose meaning all competent thinkers agree."* A brief review of the number of different intelligence tests will clearly indicate serious lack of agreement among the mental testers themselves. In fact, psychologists offer no authoritative or even quasi-unanimous definition of "intelligence". How, then, may we hope for an objective standard of measurement, when the precise specifications of the commodity to be measured are still lacking?

When considering the objectivity and reliability of the mental test, another question inevitably presents itself: *Who has measured the intelligence of the mental tester, so as to determine his qualifications to prepare a test—that is, to create this objective standard by which the minds of other human beings are to be measured?* As matter of fact, how are the various tests pre-

---

**Measurement in Higher Education,* page 141-142.

pared? In general, questions are first framed by the self-constituted mental tester, questions which, as he believes, are calculated to test innate mental endowment. Is not this first step subjective in the extreme? A more or less numerous group of individuals is then tested by this series of questions. On the basis of the results of this initial test, the examiner draws his conclusions as to the adequacy of the questions. He may make a number of tests of this kind, modifying the questions in accordance with his findings. His findings are again subjective, because they are based upon his inferences and interpretations of results. Ultimately, there is evolved, out of his own mind, a series of questions which are answerable by a certain arbitrary percentage (say 65%) of the more or less homogeneous group under examination. That particular series of questions is then adopted as a norm or standard for that particular group and for similar groups. On the foundation of such narrow individual subjectivity, how can we hope to create an objective scale of measurement?*

Finally, where is the virtue in pure 'objectivity' when we are dealing with human values? Is it not true that, by its very nature, intelligence can be judged only by intelligence; reason can be recognized and appreciated only by reason? The "light which lighteth every man that cometh into the world" is an internal light, by its nature, 'subjective'. Only by projecting that inner light outward in rational activity do we gain insight into the mind of our fellowman. Reason recognizes itself in the reason of another. Mind unites with mind. Then 'subject' and 'object' are united. Only by such a union do we have true knowledge, based upon understanding and sympathy.

---

*It is interesting to note that the original questions, framed by Binet for French children, failed to give satisfactory results when applied to California children. Is it pertinent to inquire why the Stanford revision of the Binet test, based upon the examination of California children, should be expected to provide an adequate scale of measurement for New York children, for instance?

II. OBJECTIONS BASED UPON PRINCIPLES INVOLVED.

We now turn to a consideration of the fundamental hypothesis upon which all forms of mental tests are based—namely, the possibility of measuring human intelligence by an objective, mechanical, standardized test, in such a way as to foretell mental capacity in advance of achievement. The validity of this presupposition the critics challenge, and their position may be summarized somewhat as follows:

1. Individual Differences.

Critics of the intelligence test, of course, recognize individual differences in mental capacity and in power of achievement. They maintain, however, that the power of achievement may not be measured accurately in terms of intelligence alone. In the first place, it is impossible to 'isolate' intelligence so as to measure it; and secondly, other qualities, which are usually included in the word 'character' likewise play an important rôle in the history of any individual. The new type of examination makes no adequate provision for the test of character, and yet there is no limit, which we can prophesy, to the influence of character on mental achievement. Those qualities which go to make character are among the imponderables and immeasurables, and are as elusive as the twinkle of a star, but they are none the less real and potent in the development of the individual. Indeed, it is a platitude to remark that for lack of character, a superior intelligence may lead the individual to ruin; the power of that superior intelligence may, for lack of character, be vitiated and perverted. Likewise, the possession of character may greatly enhance the intellectual attainments of the mediocre mind.

To remedy this obvious defect, effort has been made to formulate standardized tests which wll furnish a 'scientific' basis for testing character, and will thus provide a necessary supplement to the present intelligence test. The futility of such an attempt will be apparent, when we consider that there is no limit to the possible combinations of characteristics in one individual. A standard is therefore impossible, for no 'standardized test' could match more than one individual. This infinite variety is one of the glories of the race, one of the mysteries of life, and any attempt to reduce it to precise measurement or to describe it in mechanistic or mathematical terms, seems not only the utmost folly but, when rightly understood, it seems little short of sacrilege. As John Stuart Mill says, the aim of Society should be the development of humankind in its richest diversity. If this be true, any educational instrument, which attempts to set up a norm, and to rank as "superior" those who most nearly correspond to that norm or standard, must be subversive of the best interests of society and humankind, because it sets a premium on the obliteration of individuality.

In the lower schools, standardization may have the virtue of expediency—if that be a virtue—but certainly in the higher reaches of education, the aim should be toward the encouragement of individuality, leaving as free play as possible to the expression of the "richest diversity".

Moreover, the human being is an *organic* whole, spiritually as well as physically, and in studying character we are bound to come to the conclusion that the whole is *more* than the sum of its parts. That *whole* defies the power of scientific analysis or measurement. No description of a human being is adequate, which includes only an estimate of his particular qualities, such as perseverance, initiative, integrity, and so on through the whole gamut of qualities which might be subsumed under 'character'. All these qualities must be considered together, interplaying one with another, and in relation to the whole. Only then do we approach true understanding.

You may say, we are constantly judging human beings. To be sure, the 'divine spark' within us may give us insight into the probable ability possessed by any particular individual. But in order to reach such a judgment, we must consider the individual as a whole—not only his mentality in the narrower sense of the term, but all the many other qualities which contribute to success or failure. For weakness at one point there may be compensatory strength at another point. We may begin with analysis of the particular elements which go to make up any individual, but our judgment is in no sense complete, or valid, or true, until that analysis has been transformed into a synthesis, until we understand those particular elements in relation to each other and to the whole. It is in this way that we judge our friends;

likewise, in this way, teachers must judge the capacity of their students.

The increase in numbers in our schools greatly increases the difficulty of the practical problem, especially when the time comes for transfer from one institution to another, and above all from the secondary school to the college or university. At this time the bond of human sympathy and understanding between student and teacher is temporarily lost; the new bond has not yet been formed. This seems to be one of the chief reasons for the many errors in judgment in selecting students for admission to college. Is this defect to be successfully removed by further de-humanizing the college-entrance machinery, or ought our efforts to be exerted rather in the other direction?

In judging individual differences, it is also important to remember that man is not born with reason, but only with the *capacity to reason.* The development of that capacity depends upon many things within and without the individual. That is to say, the child is only *potentially* rational. The conversion of that potentiality into actuality demands effort, persistent and determined effort, and the process may continue throughout life. To a certain degree, this innate power may develop early, as in the precocious (but often superficial) child. In other cases, this power may lie dormant for a much longer period and ripen more slowly, but slow growth may mean a finer fruit. Just as there is considerable age-variation among individuals in their physical maturity, so there is in their psychic maturity. Moreover, the child that is slow to learn often learns more thoroughly and retains what he finally learns more tenaciously than the child who catches hold more quickly and lets go more easily. "Easy come, easy go" may be true of spiritual as of material substance.

2. Minds at Their Greatest Stretch.

Similarly, critics of the intelligence test will readily grant that the mental capacity of any individual may be definitely fixed and limited by nature, and it may be that no one can add a cubit to his mental stature. So far they may agree with the Fatalism of the mental testers. But they with equal emphasis deny the power of any mortal device to measure that innate mental capacity, or to foretell the possibilities of its future development, or to judge that capacity except by results and the methods of accomplishing results. "A man never knows what he can do till he tries" is an old adage, but it is as true today as it ever was.

Few of us exercise our minds at their greatest stretch. In other words, the individual rarely, if ever, actually reaches the limits of his mental capacity. That may seem a startling assertion. It does not deny the fact of 'mental strain' in particular cases, but it suggests that mental strain does not come from using the mind, but from abusing it. In any particular direction, to be sure, an individual may face what may appear to be, and what may actually be, an insurmountable obstacle, but that does not mean that his capacity may not lie in another direction. On the other hand, an obstacle may be, and often is, an incentive or stimulus to achievement. How often do men surprise themselves and their friends by the display of unsuspected power. It may be under stress of sudden emergency, of strong emotion, or even in the most ordinary circumstances, a "light may begin to dawn" upon us, and with that light may come a corresponding and undreamed of strength.

These critics maintain that the real mind is the potential mind, the mind at its greatest stretch. How far it is possible for a particular mind to stretch no mortal man knoweth. Mental tests are based upon the assumption of a static intelligence level. This assumption is unwarranted, for mental life in its essence is pure activity—a constant becoming, not mere being. That is to say, we cannot hold the mind still long enough to measure it with any degree of accuracy. What it is today is no certain criterion of what it may be next week or next year. Especially is this true in the pre-college years. The "dull boy" of today may have a sudden awakening, and may startle his family and his teachers by rapid development.

The charge is further made that the intelligence test provides no adequate intellectual or emotional stimulus for a mind to stretch; and even if it did, the stretch at any one time would not necessarily indicate its potential stretch. So much depends upon the mood the student happens to be in at the time of the examination, upon his physical condition, upon his early training, home surroundings, upon his social and economic status, and upon a thousand other considerations which no standardized mental test can possibly make adequate allowance for.

Moreover, any particular individual has different capacities, different levels within his mentality, some near the surface, some deeper. The intelligence test calls into play only the superficial levels; it does not reach the deeper-lying levels, which can be brought into play only by reflection or through emotional stimulus, or both.

3. The Spirit of the Exact Sciences.

One aim of the mental testers seems to be to find a standard of measurement for intelligence, apparently losing sight of the fact that intelligence is not an abstraction like length and weight, which may be arbitrarily standardized and accepted, and officially recorded at the Bureau of Standards at Washington. Another aim of the mental testers seems to be to find a method which will work automatically, as a machine, with the minimum of human interference or human guidance. The instructions of the examiner are so simple and so explicit and precise, that the conduct of the examination is to be as regular and certain as clockwork. "Uniformity" and "standardization" are the watchwords, just as they are in the chemical or physical laboratory. The questions, prepared in accordance with the specifications of a standard model, have been printed by the wholesale, and adapted to large groups. The numerical grades can be computed "almost mechanically by an average clerical helper under the supervision of a trained examiner."*

The result, obviously, is a great saving in time and money, as well as the elimination of the 'personal equation'. To the extent that this aim is attained, may we say that the examination-procedure has been mechanized, and to a corresponding degree, de-humanized?

A similar aim has in the past been pursued by 'big business'. In the sacred name of "Efficiency", there have been more or less successful eifforts on the part of the heads of industry to reduce human beings to machines and automata. Today there seems to be a happy tendency on the part of the more enlightened entrepre-

---

**Measurement in Higher Education*, page 31.

neurs to seek ways and means of re-introducing the human element into business. Is it not strange that, just at a time when the commercial world is attempting to discard this mechanical method, and is realizing the folly of any attempt to mechanize humanity and human relations and human power, the educational world should be so industriously endeavoring to adopt a similar method and to apply it to the higher levels of human life?

The critics protest against this effort to apply the categories of the exact sciences to the analysis and measurement of the human mind. The physicist can measure the potential energy of any material substance, before that potential energy has been transformed into kinetic energy. Similarly, the mental testers assure us that they can, with a reliable degree of accuracy, measure the 'potential energy', or capacity, of the individual mind, before that potentiality has become actuality. But let us not forget that the physicist and the psychologist are dealing with entirely different 'substances'. The human mind and its activities cannot truly be considered by the same categories as phenomena in the physical universe. Mental activity, the human spirit, is something unique and irreducible to physical terms. It is immanent in the physical universe, but it is in no sense identical with physical phenomena. We may say, with Aristotle, that the spiritual and the material are one and inseparable—yes—but it is equally true that they are different and distinguishable. The Stagirite will support us in this contention also. Spirit, intelligence, reason, rationality—that part of us which is embraced in the word 'humanity', cannot be expressed in quantitative terms, cannot be described by a mathematical symbol or formula. It is intangible, imponderable, immeasurable, acccording to the standards of exact science.

LUCY SHEPARD CRAWFORD,
Sweet Briar College
Virginia

INDIAN PRESS
SURBER-ARUNDALE CO.
CHARLOTTESVILLE, VA.